Információs rendszerek elméleti alapjai

Információelmélet

Az entrópia tulajdonságai

1. Az n elemű halmazon egyenletes eloszlás entrópiája, $A(n)=H(1/n, 1/n, \ldots 1/n)=log_2 n$
2. A két elemű eloszlások (p, 1-p) jellemzése

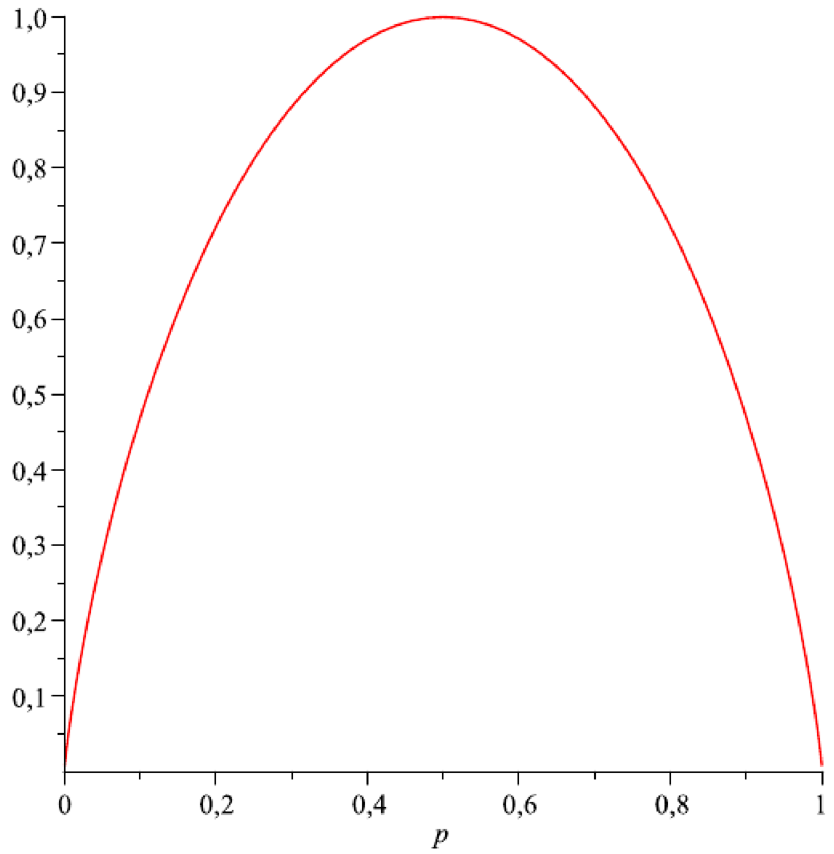
$$H(p,1-p)=-p \quad log_2-(1-p)log_2(1-p)$$

$$H(p,1-p)\leq H\left(\frac{1}{2},\frac{1}{2}\right)=1, \quad egyenlő \quad \text{csak} \quad p=\frac{1}{2}$$

$$H(p,1-p) \quad \text{szigorúan monoton növekvő}\left[0,\frac{1}{2}\right],$$

$$\text{szigorúan monoton csökkenő}\left[\frac{1}{2},1\right].$$

$$H(p,1-p)$$



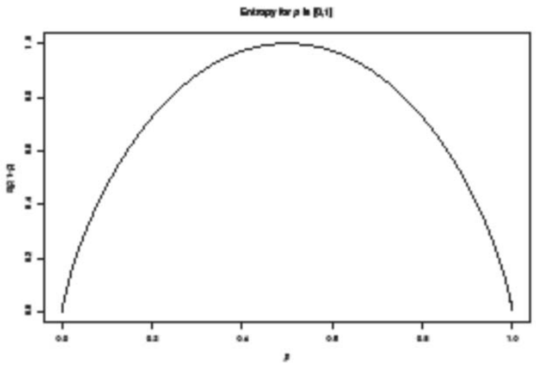


Figure 1.4

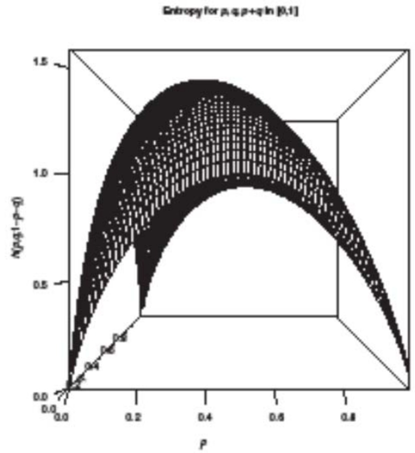


Figure 1.5

The graph of $\eta(p)$ is plotted in Figure 1.4. Observe that the graph is concave as $\frac{d^2}{dp^2}\eta(p) = -(\log e)/[p(1-p)] < 0$. See Figure 1.4.

The graph of the entropy of a three-point distribution

$$\eta_3(p,q) = -p\log p - q\log q - (1-p-q)\log(1-p-q)$$

Az entrópia tulajdonságai

3. $$H\left(p_1,\ldots,p_n\right) \leq H\left(\frac{1}{n},\ldots,\frac{1}{n}\right) = A(n) = log_2 n$$

Egyenlőség csak $p_i=1/n,\ i=1,\ldots,n$ teljesül.

Bizonyítás: Tegyük fel, hogy $\exists \quad p_{n-1} \neq p_n$.

Az elágazási szabály

$$H(p_1,\ \ldots\ p_{n-2}, p_{n-1}, p_n) =$$

$$H(p_1, p_2,\ \ldots, p_{n-1} + p_n) + (p_{n-1} + p_n)\, H\left(\frac{p_{n-1}}{p_{n-1} + p_n}, \frac{p_n}{p_{n-1} + p_n}\right)$$

$$\leq H(p_1, p_2,\ \ldots, p_{n-1} + p_n) + (p_{n-1} + p_n)\, H\left(\frac{1}{2}, \frac{1}{2}\right)$$

Az eloszlás kiegyenlítése felé történő változás, (elmozdulás az egyenletes eloszlás felé), növeli az entrópiát

Az entrópia tulajdonságai

4. Valószínűségi változók entrópiája:

Definíció: valószínűség-eloszlások entrópiája

$$P(\xi = x_i), \quad i = 1,\dots,n, \quad \xi \text{ eloszlása } \sum_{i=1}^{n} P(\xi = x_i) = 1$$

$$P(\eta = y_j),\ j = 1,\dots,m, \eta \text{ eloszlása } \sum_{i=1}^{n} P(\eta = y_i) = 1$$

$$(\xi, \eta)\ \textbf{\textit{együttes eloszlása}},\ p_{ij} = P(\xi = x_i, \eta = y_j).$$

A teljes valószínűség tétele alapján:

$$P(\xi = x_i) = \sum_{i=1}^{m} P(\xi = x_i, \eta = y_j), i = 1,\dots,n$$

$$P(\eta = y_j) = \sum_{i=1}^{n} P(\xi = x_i, \eta = y_j), j = 1,\dots,m$$

$$P(\xi = x_i, \eta = y_j) = P(\eta = y_j) P(\xi = x_i | \eta = y_j),$$

A ξ feltételes eloszlása, $\eta = y_j$ feltétel szerint,

$$P(\xi = x_i | \eta = y_j) = \frac{P(\xi = x_i, \eta = y_j)}{P(\eta = y_j)} = \frac{P(\xi = x_i, \eta = y_j)}{\sum_{i=1}^{n} P(\xi = x_i, \eta = y_j)}$$

Az entrópia tulajdonságai

Az előbbiek szerint az egyes eloszlások entrópiáját véve

$$H(\xi) = -\sum_{i=1}^{n} P(\xi = x_i) log_2 P(\xi = x_i), i = 1, \dots, n$$

$$H(\xi, \eta) = -\sum_{i=1}^{n} \sum_{j=1}^{m} P(\xi = x_i, \eta = y_j) log_2 P(\xi = x_i, \eta = y_j),$$

$$H(\xi, \eta = y_j) = -\sum_{i=1}^{n} P(\xi = x_i | \eta = y_j) log_2 P(\xi = x_i | \eta = y_j)$$

Az entrópia tulajdonságai

1. Tétel: A független valószínűségi változók együttes entrópiája az egyes változók entrópiájának összege:

<u>Bizonyítás</u>: Ha η, ξ val. változók. Függetlenek, akkor

$$P(\xi = x_i, \eta = y_j) = P(\xi = x_i)P(\eta = y_i)$$

$$H(\xi,\eta) = -\sum_{i=1}^{n}\sum_{j=1}^{m} P(\xi = x_i)P(\eta = y_i)[log_2 P(\xi = x_i) + log_2 P(\eta = y_i)] =$$

$$-\sum_{i=1}^{n} P(\xi = x_i)\underbrace{\sum_{j=1}^{m} P(\eta = y_i) log_2 P(\eta = y_i)}_{H(\eta)} - \sum_{j=1}^{m} P(\eta = y_i)\underbrace{\sum_{i=1}^{n} P(\xi = x_i) log_2 P(\xi = x_i)}_{H(\xi)} =$$

$$H(\eta) + H(\xi)$$

Az entrópia tulajdonságai
2.Tétel:

2.Tétel: Az egyenlőség a függetlenség szükséges és elégséges feltétele :

$$H(\xi,\eta) \le H(\eta) + H(\xi)$$

Az egyenlőség akkor és csak akkor teljesül, ha ξ, η függetlenek.

(Az entrópia additivitása a függetlenség karakterisztikus tulajdonsága)

Bizonyítás: a feltételes entrópia egy általánosabb tulajdonságából következik. $H(\xi|\eta)$.

Kérdés: Mennyi bizonytalanság marad átlagosan ξ-re nézve, ha lehetőségünk lesz η előzetes megfigyelésére. A teljes bizonytalanságból $H(\xi,\eta)$, $H(\eta)$ ismert, tisztázott; ξ , megfigyelése előtt, tehát marad $H(\xi,\eta) - H(\eta)$

Az entrópia tulajdonságai

2.Tétel:

Másként megfogalmazva,

Ha az elágazási szabály alapján – η megfigyeléséig visszük az elágazást – az elágazás utáni részt tekintjük *ξ-nek,* az η megfigyelése után megmaradó átlagos bizonytalanságának.

Az elágazási szabály szerint

$$H(\xi,\eta) = H(\eta) + \sum_{j=1}^{m} P(\eta = y_j) H(\xi | \eta = y_j)$$

Tehát mindkét úton ugyanazt kapjuk.

Definició: A ξ valószínűségi változó η szerinti feltételes entrópiája,

$$H(\xi|\eta) = H(\xi,\eta) - H(\eta) = \sum_{j=1}^{m} P(\eta = y_j) H(\xi | \eta = y_j)$$

Az entrópia tulajdonságai
2.Tétel:

A $H(\xi|\ \eta=y_j)$ kifejtve

$$(6)\quad H(\xi|\eta)=-\sum_{j=1}^{m}P(\eta=y_j)\sum_{i=1}^{n}\frac{P(\xi=x_i,\eta=y_j)}{P(\eta=y_j)}log_2\frac{P(\xi=x_i,\eta=y_j)}{P(\eta=y_j)}=$$

$$H(\xi|\eta)=+\sum_{j=1}^{m}\sum_{i=1}^{n}P(\xi=x_i,\eta=y_j)log_2\frac{P(\eta=y_j)}{P(\xi=x_i,\eta=y_j)}$$

$\langle$A fenti formulát használjuk a lemma bizonyítására$\rangle$

$$P(\xi=x_i|\eta=y_j)=\frac{P(\xi=x_i,\eta=y_j)}{P(\eta=y_j)}$$

A 2. tétel átírása a feltételes entrópiára nézve, az egyenlőtlenség mindkét oldalából *H(η)* levonásával

$$H(\xi,\eta)\le H(\eta)+H(\xi)\Leftrightarrow H(\xi|\eta)\le H(\xi)$$

$$2'.\text{tétel}: H(\xi|\eta)\le H(\xi)\ egyenl\H{o}\ \Leftrightarrow\quad ha\quad \xi\quad \text{és}\quad \eta\quad \text{függetlenek}$$

A 2'. Általánosabb egyenlőtlenséget bizonyítunk

3. Tétel

Az entrópia tulajdonságai
3. Tétel

A 2'tételnél általánosabb egyenlőtlenséget bizonyítunk

3. Tétel

$H(\xi|\eta) \leq H(\xi|f(\eta))$, és az egyenlő

akkor és csak akkor teljesül,

ha $\forall\ x,y,z$ -re , ha $f(y) = z$ -re fennáll

$$P(\xi = x|\eta = y) = P(\xi = x|f(\eta) = z)$$

Az entrópia tulajdonságai
3. Tétel

Bizonyítás: szükségünk van f(η), (ξ, f(η)) eloszlására.

$$P\big(f(\eta)=z\big)=\sum_{f(y)=z}P(\eta=y),$$

$$P\big(\xi=x,f(\eta)=z\big)=\sum_{f(y)=z}P(\xi=x,\eta=y)$$

Az entrópia tulajdonságai
3. Tétel

Írjuk fel *H(ξ|f(η))* feltételes entrópiát *(6)* szerint

$$H(\xi|f(\eta)) = \sum_x \sum_z P(\xi = x, f(\eta) = z) log_2 \frac{P(f(\eta) = z)}{P(\xi = x, f(\eta) = z)} =$$

$$\sum_x \sum_z \left(\sum_{f(y)=z} P(\xi = x, \eta = y) log_2 \frac{\sum_{f(y)=z} P(\eta = y)}{\sum_{f(y)=z} P(\xi = x, \eta = y)} \right)$$

Az entrópia tulajdonságai
3. Tétel

Hasonlóan *H(ξ|η)* felírását végezzük *x,z, f(y)=z,* bontásban

$$H(\xi|\eta)=\sum_x\sum_z\left(\sum_{f(y)=z}P(\xi=x,\eta=y)log_2\frac{\sum_{f(y)=z}P(\eta=y)}{\sum_{f(y)=z}P(\xi=x,\eta=y)}\right)$$

Elegendő az x,z összegzésen belüli tagokat összehasonlítani. Az összehasonlítandó formulák alakja?

$a_1,a_2,\ldots, a_n>0$, $b_1, b_2, \ldots, b_n>0$, $a=\Sigma_{i=1}^n a_i$, $b=\Sigma_{i=1}^n b_i$

$$\sum_{i=1}^n a_i\, log_2\frac{b_i}{a_i} \quad \text{és} \quad a\, log_2\frac{b}{a},$$

$$a_i \cdots P(\xi=x,\eta=y), b_i \cdots P(\eta=y)$$

$$a=\sum_{f(y)=z}P(\xi=x,\eta=y), b=\sum_{f(y)=z}P(\eta=y)$$

Az entrópia tulajdonságai
3. Tétel - Lemma

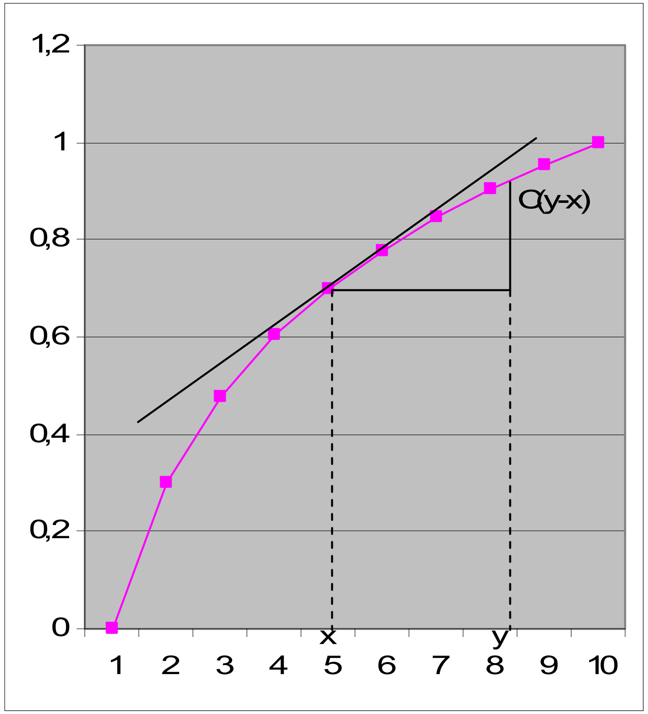
Lemma:

$$\sum_{i=1}^{n} a_i \log_2 \frac{b_i}{a_i} \le a \log_2 \frac{b}{a}, \quad a = \sum_{i=1}^{n} a_i, b = \sum_{i=1}^{n} b_i$$

Egyenlőség csak $b_i/a_i=b/a$ esetén, áll fenn

Bizonyítás: a $\log_2 x$ függvény konvexitása miatt

Közelítés



Az entrópia tulajdonságai

3. Tétel - Lemma

$$log_2\, y \le log_2\, x + C_x\left(y - x\right)\cdots \text{és egyenlő csak } x = y$$

$\rightarrow$

$$\log_2 \frac{b_i}{a_i} \le log_2 \frac{b}{a} + C_{\frac{b}{a}}\left(\frac{b_i}{a_i} - \frac{b}{a}\right), \text{és egyenlő csak } \frac{b_i}{a_i} = \frac{b}{a},$$

szorozva $\quad a_i - vel \quad$ összegezve

$$\sum_{i=1}^{n} a_i\, log_2 \frac{b_i}{a_i} \le \sum_{i=1}^{n} a_i\, log_2 \frac{b}{a} + C_{\frac{b}{a}}\left(\sum_{i=1}^{n} a_i\left(\frac{b_i}{a_i} - \frac{b}{a}\right)\right) =$$

$$a\, log_2 \frac{b}{a} + C_{\frac{b}{a}}\left(\underbrace{\sum_{i=1}^{n} b_i}_{b} - \underbrace{\sum_{i=1}^{n} a_i}_{a} \frac{b}{a}\right) = a\, log_2 \frac{b}{a}$$

Az entrópia tulajdonságai

3. Tétel

Lemma miatt

$$H(\xi|\eta) \leq H(\xi|f(\eta))$$

És egyenlőség csak akkor. ha $\forall$ *z*-re teljesül a lemma azaz *f(y)=z-n*

$$\frac{P(\eta = y)}{P(\xi = x, \eta = y)} = \frac{\sum_{f(y)=z} P(\eta = y)}{\sum_{f(y)=z} P(\xi = x, \eta = y)}$$

Az entrópia tulajdonságai
3. Tétel-2. Tétel

3. Tétel következménye a 2. tétel az f(y) $\equiv$ c függvényre

És egyenlőség csak akkor. ha $\forall$ *z*-re teljesül a lemma azaz *f(y)=z-n*

$$H(\xi|\eta) \leq H(\xi|f(\eta)) = P(f(\eta)=c)H(\xi|f(\eta)=c) = H(\xi)$$

$$= \quad \Leftrightarrow$$

$$\frac{P(\xi=x,\eta=y)}{P(\eta=y)} = \frac{P(\xi=x, f(\eta)=c)}{P(f(\eta=c))} = P(\xi=c)$$

Tehát függetlenek (az egyenlőség szükséges és elégséges feltétele.

2’. Tétel :

$$H(\xi_3|\xi_1,\xi_2)\le H(\xi_3|\xi_2)$$

$$= \quad \Leftrightarrow \quad ha \quad P(\xi_3=x_3|\xi_1=x_1,\xi_2=x_2)=P(\xi_3=x_3|\xi_2=x_2) \quad \forall \quad x_1,x_2,x_3$$

Bizonyítás :

$$f(x_1,x_2)=x_2 \quad \text{választással}$$

$$\mathrm{H}\left(\underbrace{\xi_3}_{jöv\H{o}}\middle|\underbrace{\xi_1\ ,\ \overbrace{\xi_2}^{jelen}}_{múlt}\right)\le H(\xi_3|f(\xi_1,\xi_2))=H(\xi_3|\xi_2)$$

$$= \quad \Leftrightarrow \quad ha$$

$$P(\xi_3=x_3|\xi_1=x_1,\xi_2=x_2)=P(\xi_3=x_3|f(\xi_1,\xi_2)=x_2)=$$

$$P(\xi_3=x_3|\xi_2=x_2)$$

Az entrópia tulajdonságai

2’. Tétel Általánosan megfogalmazva

Markov lánc tulajdonság

$$H\left(\xi_t = x_t \middle| \xi_1 = x_1, \ldots, \xi_{t-1} = x_{t-1}\right) \leq H\left(\xi_t = x_t \middle| \xi_{t-1} = x_{t-1}\right)$$

$$= \quad \Leftrightarrow \quad ha \quad P\left(\xi_t = x_t \middle| \xi_1 = x_1, \ldots, \xi_{t-1} = x_{t-1}\right) = P\left(\xi_3 = x_3 \middle| \xi_{t-1} = x_{t-1}\right)$$

$$\forall \quad x_1, \ldots, x_{t-1}, x_t$$

Az entrópia tulajdonságai
2'. Tétel

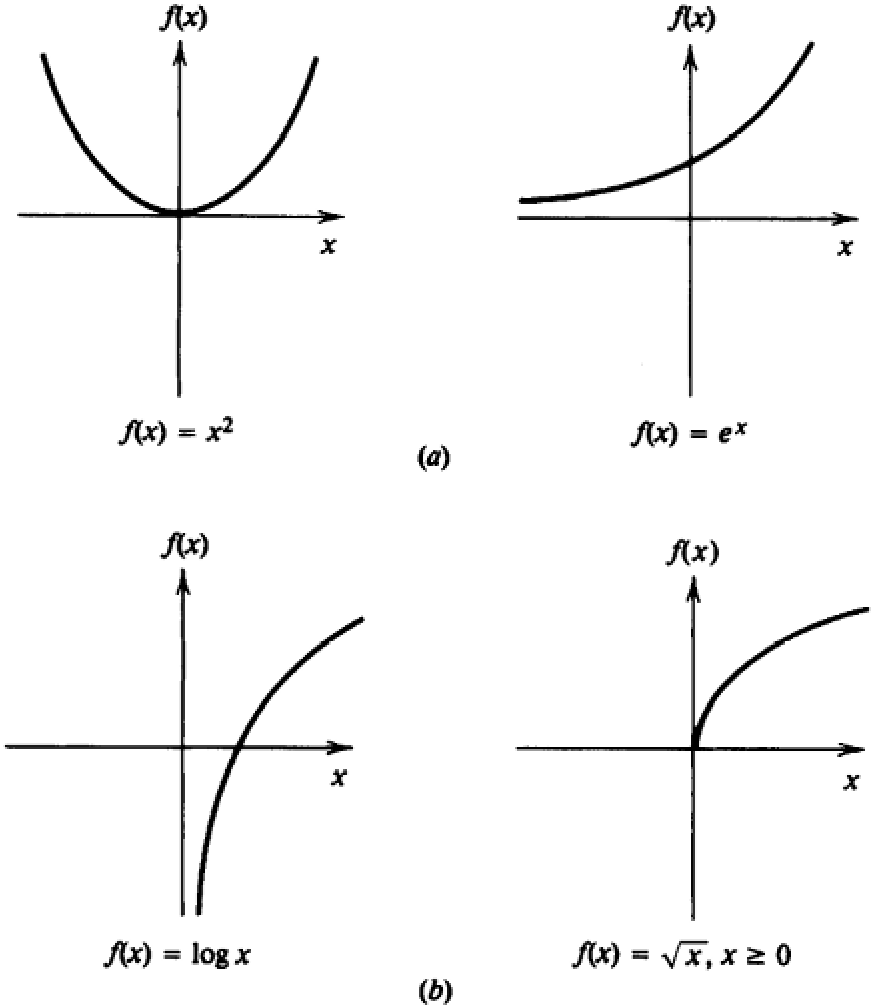
2'. Tétel :

„ a legnagyobb a bizonytalanság a jövőre, akkor van ha csak a jelent ismerem és a múltat nem”

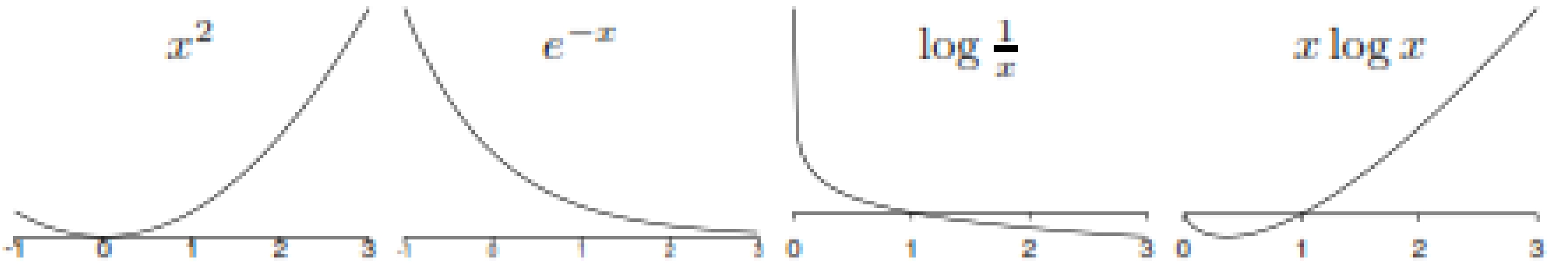
Markov lánc tulajdonság: „a jelen ismeretében a múlt és a jövő független”

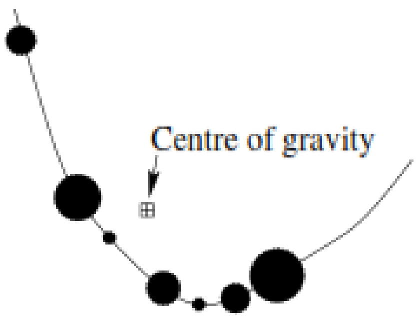
„A jövő a múlttól csak a jelenen keresztül függ”

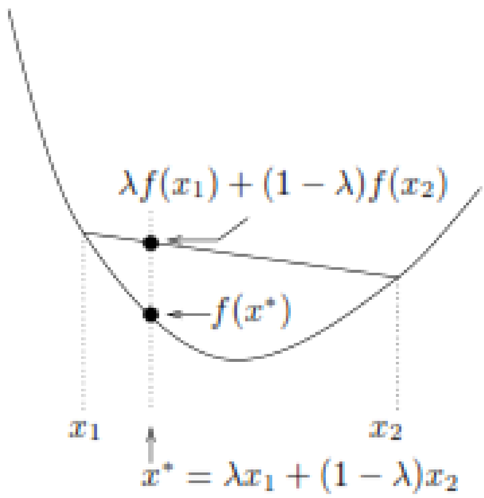
Konkáv és konvex függvények



konvex függvények







Konvexitás definiciója

Matematikai illusztráció
A lemma következményei

$$(1)\quad \sum_{i=1}^{n} a_i \log_2 \frac{b_i}{a_i} \le a \log_2 \frac{b}{a}$$

$$\sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} q_i = 1 \ \forall i-re, p_i, q_i \ge 0$$

$$\sum_{i=1}^{n} p_i \log_2 \frac{p_i}{q_i} \le 1 \log_2 \frac{1}{1} = 0$$

$$-\left(\sum_{i=1}^{n} p_i \log_2 p_i\right) - \left(-\sum_{i=1}^{n} p_i \log_2 q_i\right) \le 0$$

$$H(p_1,\ldots,p_n) \le -\sum_{i=1}^{n} p_i \log_2 q_i$$

$$(2)\ a_i = p_i,\ b_i = 1, \quad b = n$$

$$\sum_{i=1}^{n} p_i \log_2 p_i^{-1} \le 1 \log_2 \frac{n}{1} = \log_2 n$$

Az egyenletes eloszlás entrópiája maximális

Matematikai kitérő – A valószínűségszámítás alapfogalmairól

Elméleti összefoglaló

A $B_1, B_2, \ldots, B_n$ események teljes eseményrendszert alkotnak, ha $B_1 + B_2 + \ldots + B_n = \Omega$ és $B_i \cdot B_j = \emptyset$, ha $i \neq j$ (i=1, 2, ... ,n; j=1, 2, ..., n) .

[Másképpen $P(B_1)+\ldots+P(B_n)=1$ és $P(B_i \cdot B_j)=0$]

A teljes valószínűség tétele

Ha a $B_1, B_2, \ldots, B_n$ események teljes eseményrendszert alkotnak, és $P(B_i) \neq 0$ (i=1, 2, ... ,n) , akkor tetszőleges A esemény valószínűségére érvényes a következő:

$P(A)=P(A|B_1)\cdot P(B_1)+P(A|B_2)\cdot P(B_2)+\ldots+P(A|B_n)\cdot P(B_n)= \sum_{i=1..n} P(A|B_i)\cdot P(B_i)$.

A Bayes-tétel

Ha a $B_1, B_2, \ldots, B_n$ események teljes eseményrendszert alkotnak, és $P(B_i) \neq 0$ (i=1, 2, ... ,n) , továbbá A tetszőleges olyan esemény, amelyre $P(A) \neq 0$, akkor

$P(B_i|A)= P(A|B_i)\cdot P(B_i) / \sum_{j=1\ldots n} P(A|B_j)\cdot P(B_j)$.